DOCUMENT RESUME

ED 196 957 TM 810 109

AUTHOR Chapman, Paul Davis
TITLE 'A Mental Test for Every Child': The Use of Intelligence Tests in Progressive School Reform.
PUB DATE 10 Apr 80
NOTE 40p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Ability Grouping; Educational Change; *Educational History; Educational Trends; *Intelligence Tests; *Progressive Education; *Psychological Testing; Testing Problems
IDENTIFIERS Test Use

ABSTRACT

This essay shows that the adoption of intelligence tests by the schools was a complex development. Tests were adopted during the 1920s as part of the reform program fashioned by the network of applied psychologists and school people. While the network itself often viewed testing as a means to improve the schools and society, immigrants and blacks did not share their vision. The adoption of the tests proceeded quickly for several reasons: (1) the need for an efficient means of sorting students; (2) the catalyst of World War I; and (3) the consensus among school people that the use of tests promoted a more efficient and scientific education. In the last two decades the progressive reform of intelligence testing and ability grouping has come under sharp attack. Revisionist historians and social activists have criticized intelligence tests as racially biased, elitist, and as obstacles to social equality. Historically it cannot be denied that the rhetoric and the programs of progressive reformers occasionally led to inequality in the schools. Yet as the nation seeks to reevaluate policy regarding the use of tests, it is essential to consider the origins of intelligence testing in the progressive era. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED196957

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

'A Mental Test for Every Child':
The Use of Intelligence Tests in Progressive School Reform

Paul Davis Chapman
San Francisco University High School
San Francisco, California

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. D. Chapman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper Presented at the Annual Meeting of the
American Educational Research Association, Boston
10 April 1980

Not for Quotation or Citation
Without Permission of the Author

TM 810 109

"'A mental test for every child' is no longer an unreasonable slogan declared Lewis Terman in 1920. As a leader in the intelligence testing movement during the progressive era, Terman spoke with authority. He also spoke with vision, for by 1925 millions of intelligence tests were being given annually to school children and increasingly the tests were being used as a means of sorting students into ability groups.[1]

The rapid adoption of intelligence tests in schools has been the subject of several recent studies of progressive reform. Revisionist interpretations have argued that widespread testing emerged in an atmosphere of conflict and that the tests were foisted upon the schools by groups seeking to promote their own social philosophy or to protect their own self-interest. More traditional accounts have stressed that the tests were adopted in a spirit of consensus and as a great contribution of applied psychology to social and educational progress. The origin of intelligence testing continues to be an important subject, not only because of the current debate about the use of tests in schools, but also because the history of testing can help illuminate the nature of progressive educational reform. This essay suggests several reasons why intelligence tests were developed and used in American schools.[2]

In one sense the testing movement was a top-down reform, conceived by a network of newly emergent professionals--university professors of psychology and public school administrators--who advocated the universal use of tests to classify students. This group was supported by philanthropic foundations, educational publishers, and such national institutions as the Bureau of Education. The tests they developed reflected many progressive values such as efficiencey, conservation, a belief in science

and the professional expert, a concern with individual differences, and a nativistic fear of the new immigrants. As a result of this consonance between certain progressive values and the tests, the adoption of IQ tests was hastened.

Recent evaluations of the testing movement have focused on the network of applied psychologists responsible for developing and marketing the tests. An element missing in this analysis has been the perspective of school administrators and the problems they faced at the turn of the century. This essay argues that in addition to the "push" from the network, intelligence tests were adopted as a result of the changing structure and function of the schools. At the turn of the century, school people faced a vast array of problems. Urban school enrollments mushroomed as a consequence of massive immigration, the move from the country to the city, and newly enforced compulsory education laws. School populations represented, as well, an increasingly diverse patchwork of ethnic groups. At roughly the same time, schools evolved a new set of objectives, and the school curriculum became much more differentiated. The common school curriculum and classical studies of the nineteenth century gave way to preparation for life. By the 1920s schools offered courses of study ranging from the liberal arts to manual and "home" arts and included commercial and vocational programs. School administrators needed a means of classifying students in order to assign them to these different curricular tracks. Intelligence tests provided such a means.

In an earlier study, "Schools as Sorters: Testing and Tracking in California, 1910-1925," I examined the adoption of tests in three California case studies. This essay explores the origins of intelligence testing in a national context. To explain how and why the intelligence test was conceived as an instrument of progressive reform, the paper traces the antecedents of testing in Europe and America and sketches the rise of

of a measurement network in education. Then an evaluation is made of the problems schools faced and the historical ways school people classified students to provide for individual differences. Finally, survey research from the 1920s documents the success of progressive reform--the rapid adoption of intelligence tests and their use to classify students into ability groups.[3]

The Network

The idea of measuring intelligence and using tests to classify school children began not in America but in Europe. In 1924 psychologist Kimball Young wrote a brief history of mental testing and noted: "While American psychologists have played an important role in the development of mental testing, the cultural borrowing in this instance is so obvious that we may not escape its significance." According to Young, there were at least three European contributions to the development of mental measurement: the rise of experimental psychology in Germany, the evaluation of individual differences in England, and the creation of a scale for measuring intelligence in France.[4]

Initial efforts at studying mental differences began in Germany where Wilhelm Wundt founded the first formal laboratory of experimental psychology in 1879. Wundt's investigation of human consciousness and his rigorous experimental approach attracted several Americans who were later to figure importantly in the testing movement in the United States, among them J. McKeen Cattell and G. Stanley Hall. At about the same time in England, Francis Galton was exploring individual mental differences. Following the publication of On the Origin of Species in 1859 by his cousin, Charles Darwin, Galton became interested in the inheritance of mental differences. Galton's pioneering work in statistics, as well as the opening of his own psychological laboratory in London in 1884,

stimulated interest in the tests and led to the idea that a common intelligence factor, g, was strongly influenced by heredity.[5]

It was the work of Alfred Binet in France that had the most direct influence on the development of intelligence tests in America. A member of the laboratory of physiological psychology at the Sorbonne, Binet began a systematic study of individual differences among school children in Paris in the 1890s. In 1904 the Minister of Public Instruction appointed Binet to study ways of identifying children in the Paris schools who were "subnormal," those who were not making adequate progress, so that they could be "segregated" into separate classes. By the next year Binet had created the first scale for measuring intelligence, a series of thirty mental tests of increasing difficulty, and he soon brought out a revised scale that could be used to test not merely "subnormal" children, but the "normal" and "gifted."[6]

The European investigation into intelligence was soon translated to the United States, where it took a decidedly practical turn. Cattell, one of the first Americans to study with Wundt, established a psychological laboratory in America and coined the term "mental test." In the 1890s he moved to Columbia University, which soon became the center of American test development. Cattell stimulated the movement by founding several journals of importance for the young profession of applied psychology--Science and The Psychological Review--and by forging a graduate program that would produce several leaders in the testing field after the turn of the century. Edward L. Thorndike, probably the most influential graduate, received his degree in 1898 and soon published key studies on educational psychology (1930), the theory of mental and social measurements (1904), and animal intelligence (1911).

Another major center of interest in intelligence testing was Clark University in Worcester, Massachusetts. G. Stanley Hall, the president of Clark, had also studied with Wundt at Leipzig. He helped launch the fledgling profession of psychology by founding the American Journal of Psychology, by organizing the American Psychological Association, and by sponsoring the graduate studies of several men who were to become key figures in the field. Around the turn of the century, several men studied at Clark who later played central roles in developing intelligence tests: Henry H. Goddard, J.E.W. Wallin, and Lewis M. Terman.

Of these men, Terman would play the largest role in forging a testing network and in promoting the use of tests in schools. Appointed to a position at Stanford in the new Department of Education in 1910, he began in earnest a study of intelligence tests. Experimenting with 400 school children near Stanford, Terman began a revision of Binet's scale.

Based on his initial research, Terman made some preliminary observations about the possible use of intelligence tests in 1911. He saw immediate application for the tests in the diagnosis of individual problems, for they offered a "more reliable and more enlightening estimate of the child's intelligence than most teachers can offer after a year of daily contact in the schoolroom." He then referred to recent studies of school "retardation"--the discovery by Leonard Ayres and others that a high percentage of students were overage for their grade level, either because they were making slow progress or because they had started late. "Before many years," he predicted, "it will probably become a matter of course to apply serial mental tests in the public schools to all pupils who are retarded or about to become retarded, or who give indication of unusual ability." He suggested that the tests could be used for "the scientific management of special classes for atypical children in the

public schools," responding to the quest for efficiency that had surged earlier that year with the popularization of the ideas of Frederick W. Taylor by muckraker Ray Stannard Baker. Terman concluded: "Now that the individual treatment of pupils in the schools has begun, there is no stopping short of this idea. Tests must be developed which will enable us to differentiate all degrees of intellectual ability and all kinds of intellective unevenness."[7]

In 1915 Terman's revision of Binet's intelligence test was complete and in 1916 he published a handbook for its administration, The Measurement of Intelligence. Terman's intelligence test, which he named the Stanford-Binet, was to become the most widely used individual test over the next two decades. The Stanford-Binet was criticized, however, on several grounds. Some questioned the validity of this test, its constancy, and the degree to which it was influenced by heredity. Critics argued, too, that the sample for the Stanford-Binet norms produced standards that were biased against the lower classes, immigrants and blacks. Others objected to the great expectations for the use of tests, suggesting that the tests were fallible and should not be used to determine the fate of individuals.[8]

Undaunted by this criticism, Terman began to establish at Stanford a center from which the testing movement on the West Coast would emanate. The Department of Education developed a full program to prepare teachers and administrators in intelligence testing, offering courses on retardation, the education of "defectives," the social aspects of mental deficiency, and the psychology of "exceptional" children. Terman's work in testing began to attract independent financial support; a grant from the Buckel Foundation, for example, combined with university support to provide $1,000 a year for the "psychological and pedagogical study of backward and mentally defective children." Terman and Ellwood P. Cubberley, the head of the department, planned to offer Stanford's consulting services

to "cooperate with the school systems of the state in the formation of special classes for backward or otherwise exceptional children."[9]

In the years before the war, Terman campaigned to spread the applications of his work on testing to a wider audience in conjunction with the movements for social efficiency and eugenics. Perhaps because of his own recurring bouts with tuberculosis, Terman devoted considerable energy to the issues of school health and hygiene, publishing two major works on the subject: The Hygiene of the School Child, a text for teachers, and Health Work in the Schools. At this same time Terman's services were enlisted by California governor Hiram Johnson, one of the nation's leading progressives. Johnson consulted Terman on how to increase efficiency in the state's welfare institutions and commissioned him to conduct a series of mental surveys, including a study of the Whittier State School for Boys, San Quentin Prison, an orphanage, and the study of all schools in Merced county.[10]

Terman also lent his support to the fast growing eugenics movement, which hoped to improve society by controlling the reproduction of various "defective" individuals and by restricting immigration, primarily from Southern and Eastern Europe. From his work with intelligence tests, Terman had reached several conclusions consistent with the eugenics program. Increasingly, he believed that intelligence was in good part inherited. As a result, he said, "we are rapidly becoming conscious of hitherto unsuspected power to shape human destinies, and are no longer willing to remain the passive plaything of uncontrolled social and material forces." Noting that "the evolution concept is doing its work," Terman thought the logical next step was to develop ways to control society's collective intelligence.[11]

Terman came to believe that social deviance was often the result of low intelligence and that the very foundation of civilization was threatened

by what he called the "menace of the feeble-minded." He concluded that intelligence tests could help cure poverty, crime, and delinquency by identifying "feebleminded" individuals. In a survey of schools in San Luis Obispo, California, for example, he noted that feebleminded persons constituted "probably one-fourth of our habitual criminals and a still larger proportion of our paupers and 'white slaves.'" "The children of subnormal mental endowments undoubtedly constitute one of our greatest social problems," he continued. In a report with Ellwood P. Cubberley to Stanford alumni, Terman outlined his ideas for a genetic solution to these problems:

> ...it is clear that society has few tasks more important than that of identifying the feeble-minded and providing for their institutional care. There is a growing conviction that society, in self-defense, will be driven to provide institutional care for every feeble-minded individual throughout the reproductive period.

Terman's strong language and startling views were **not**, it should be noted, unlike those of other progressive reformers. In their quest for an ordered society, they sometimes infringed upon the rights of those who were without power and were not part of the dominant culture.[12]

Throughout the years before the war, Terman played a key role in shaping the early testing movement and in developing a testing network. The experience of Terman at Stanford was characteristic of applied psychologists at other institutions across the country. Several universities figured importantly in the developing network, notably Columbia, Cornell, Harvard, and Yale in the East, and Illinois, Minnesota, Indiana, Michigan, Ohio State, and Chicago in the Midwest. The tests Terman and others created reflected many of the prevailing currents of the period--the quest for efficiency, the role of the expert, the persuasion of Social Darwinism, and the appeal of nativism. Terman's innovation also meshed will with a heightened interest in measurement throughout the educational community.

The invention of intelligence tests coincided with a broader interest in the scientific measurement of educational products and the creation of standardized achievement tests, measures of student learning as opposed to student ability. The measurement movement gained momentum after the turn of the century with the rise of achievement testing, the advent of the school survey, and the institutionalization of educational testing in school and municipal research bureaus. The establishment of this institutional network concerned with educational measurement eventually paved the way for the widespread adoption of intelligence tests.

The publication of a series of articles in 1892 by Joseph Mayer Rice undoubtedly helped to trigger the quickened interest in standardized objective testing of student achievement. Rice surveyed schools in thirty-six cities to determine "the causes of the marked variation in the general degree of excellence of the schools of various localities," concluding that the "obstacles to educational reform" lay in the lack of any objective standards by which effective teaching and productive learning could be measured. He soon set out to develop objective tests in four areas--spelling, penmanship, English composition, and arithmetic. Between 1894 and 1896 Rice examined some 100,000 school children in districts across the country, reporting his findings in The Forum. After the turn of the century, he called for a twofold educational reform based on specific standards to measure student learning and the time needed to accomplish assigned tasks Rice's articles reached a wide audience and his studies sparked universities to develop standardized tests of achievement.[13]

Thorndike and his students at Teachers College, Columbia put Rice's ideas to work. The Stone Arithmetic Test, published in 1908, was one of the first standardized achievement tests. The next year Thorndike presented the first scale to measure achievement in handwriting to the

annual meeting of the American Association for the Advancement of Science. To those interested in scientific measurement, Thorndike produced not only the instruments but a sense of purpose and conviction as well. His motto, "Whatever exists at all exists in some amount," served as the credo of those trying to measure educational products.[14]

The next ten years brought an extraordinary outpouring of educational achievement tests. In 1917 Walter S. Monroe reported to the annual meeting of the National Society for the Study of Education (NSSE) that there were 218 tests available for use--168 in elementary schools and 50 in high schools. This roster was only a sample of the instruments he catalogued in Educational Tests and Measurement that same year. Another NSSE member, Edna Bryer of the Russell Sage Foundation in New York, presented a bibliography of over six hundred articles on testing appearing in 33 journals between 1907 and 1917. Her study showed that by the war, standard tests existed for sixteen subjects ranging from algebra, English composition, and geography to history, physics, and spelling. The educational measurement movement had come of age rapidly.[15]

The focus of measurement, however, was not exclusively student achievement. In the decade before the war, the leaders of the school survey movement examined and quantified virtually every aspect of education from teaching and salaries to the quality of school buildings. University professors frequently conducted the investigations; some of the most active surveyors included Cubberley and Sears at Stanford, Charles Judd and Franklin Bobbitt at the University of Chicago, and George Strayer at Teachers College. Between 1910 and 1925 nearly two hundred cities commissioned surveys of their school systems, some sixty-seven in the period from 1910 to 1919. The surveys were of key importance for the emerging measurement movement because they frequently

employed standardized achievement tests to determine the educational quality of the school population.[16]

The survey movement stimulated some school systems to take on the task of measuring student achievement themselves; they created educational research bureaus to handle the task. As Eugene Nifenecker, Assistant Director of the New York City Bureau of Research and Reference said in 1918, "The city bureau of educational research is the direct and logical outcome of the combination of the survey movement and the movement for the use of measurement." By the outbreak of the war, at least twenty-three research bureaus had been founded across the country in cities ranging in size from New York, Chicago and Philadelphia to St. Paul, Topeka, and Des Moines.[17]

The years before the war also saw the beginning of the psychological measurement of school children using individual intelligence tests, a program that was often carried out in psychological clinics. Between 1909 and 1914 J.E.W. Wallin, Terman's colleague from Clark, conducted a series of surveys that documented the swift increase in mental and psychological testing. In a poll of 103 cities, he discovered that eighty-three (81 percent of those responding) were using psychological tests to identify feebleminded and backward children, and seventy-one cities were using the Binet tests specifically. The following table shows the reported use of tests by city size:[18]

TABLE I

EARLY USE OF PSYCHOLOGICAL TESTS, 1914

City Size	No. of Cities 1910 Census	No. of Cities Reporting Use of Psychological Tests
Over 500,000	8	8
250,000-500,000	11	11
100,000-250,000	31	15
25,000-100,000	178	26
Less than 25,000	2,034	23
	2,262	83

It is hard to generalize from Wallin's data, since his eighty-three responses represented less than 5 percent of the cities in the country and the questionnaire did not evaluate the extent of use in each city. Indeed, as Wallin observed, "while this is an extremely creditable showing, particularly in view of what it portends for the future, it is necessary to emphasize that the psychological testing in most of the cities is exceedingly meager and crude, being conducted by teachers, principals, educators, psychologists and physicians who are not specialists on the physiology, psychology and pedagogy of feebleminded, backward or other types of mentally abnormal children." Nonetheless, the number of cities reporting test use is remarkable considering that Binet had developed his intelligence scale only a few years earlier.[19]

Wallin's surveys showed that the beginning of systematic mental testing was closely associated with the development of the psychological clinic, the purpose of which was to determine "the process of first-hand observation, testing and experiment." According to Wallin there were at least four functions of psychological clinics: expert diagnosis of "mental deviation," service as a "clearing house for mentally exceptional cases," research into the nature of "mental deviation" and test development, and the dissemination of information about the nature and distribution of intelligence.[20]

By 1914 according to Wallin sixty-eight clinics had been established in a variety of institutions, as the following table shows:

TABLE II

PSYCHOLOGICAL CLINICS, 1914

Type of Institution	Number of Clinics
Public schools	19
Universities, medical schools, and normal schools	26
Hospitals for the insane	5
Institutions for the feeble-minded and epileptic	6
Juvenile courts and other courts	5
Penal and correctional institutions	6
Immigrant stations (Ellis Island)	1
	68

Wallin's data show that the clinics set up in the public schools, colleges, and universities were the most numerous. At the time of the survey, clinics were operating in nineteen school systems including Chicago, New York, Oakland, Seattle, Philadelphia, Los Angeles, and Detroit. Schoolchildren were probably given intelligence tests in the clinics operating in juvenile courts and institutions for feebleminded as well. The number of clinics in medical schools, hospitals, and institutions for the insane suggests that intelligence testing in America initially focused on the "abnormal" and pathological individuals. Only later did testing spread to include all students in schools.[21]

At the threshold of World War I, three important developments thus set the stage for the intelligence testing movement. Psychologists in Europe and America had begun a systematic and scientific study of individual differences and had discovered early ways to measure intelligence tests. At Stanford, Lewis Terman had developed the technology of intelligence tests and had fashioned a blueprint for their use in pro-

gressive reform. The educational measurement movement had produced a vast array of standardized, objective tests. Together these factors helped to create a network of psychologists ready to use intelligence tests on a larger scale. The solutions that applied psychologists created were eagerly awaited by schoolmen, who faced during this same period their own set of problems.

Classifying Students

During the years around the turn of the century, school administrators and teachers faced enormous challenges. Urban school enrollments skyrocketed because of tremendous immigration, a shift in population from the farm to the city, and newly enforced compulsory education laws. To meet the needs of an increasingly diverse community, schools developed new goals; the college preparatory curriculum of the nineteenth century was gradually replaced by one directed toward social efficiency--preparation for citizenship and the world of work, and the effective use of leisure time. Soaring costs and a campaign for efficiency put pressure on school administrators to reform the schools. In response to these problems--increasing size, diversity, rising costs, and changing objectives--the school curriculum became much more differentiated. School administrators needed a new, more efficient means of classifying students in order to assign them to different curricular tracks.

As we have seen, Lewis Terman and other applied psychologists developed intelligence tests for measuring differences in student ability during these same years. Two communities--university professors of psychology and education, research bureau directors, and heads of psychological clinics on the one hand; superintendents, principals, and teachers on the other--needed each other's assistance. Applied psychologists wanted to demonstrate the usefulness of their discipline for school

reform. School people, facing a variety of complex problems, felt that applied psychology--especially the newly developed intelligence tests--could aid their efforts. The mutual interest of these two groups thus led to the beginnings of the intelligence testing movement before the war.

The call for a different means of classifying students to provide for individual differences was not new, however. Although provision was made in separate institutions for the physically disabled in the early 1800s, at mid-century the dominant institution in public education was the common school, one that offered a uniform curriculum to students of all ages, often in a one-room school. After mid-century, though, schools began to classify students according to age. Some schools, especially those in urban areas, gradually began to incorporate into the curriculum special classes for the physically and mentally disabled, for students who were either dependent or delinquent, and for immigrants who spoke no English. Throughout the nineteenth century the structure and function of the schools evolved in response to rapid industrialization, large-scale immigration, and extensive urbanization. A new classification system to treat individual differences was part of that evolution.[22]

By World War I there had been substantial changes in the structure of the schools and their purpose. The common school had become a highly differentiated elementary school offering a variety of curricula and special classes. The high school had emerged in four distinct types--college preparatory, comprehensive, commercial, and vocational. These structural modifications reflected changes in the purpose of education. In the early nineteenth century, schools attempted to provide most

students with a minimum of education while cultivating an elite for high school and college. By World War I, educators sought to make schooling universal. A commission of leading educators in 1917 defined the purpose of high school as preparation for life.

The history of curricular differentiation provides a context for understanding those who called for educational reform in the years before World War I. These early attempts to provide for individual differences help explain the rise of ability grouping in the early twentieth century. They also help explain the widespread adoption of intelligence tests after the war.

The movement to treat individual differences in special classes and special schools, as was noted above, began early in the nineteenth century. Starting with the founding of the American School for the Deaf in Hartford in 1817, private and public institutions provided special education in the nineteenth and early twentieth century for the blind and cripples, orphans and juvenile delinquents, as well as the feeble-minded, and epileptic. Gradually city school systems added special classes for these students and then began to provide special classes for non-English speaking immigrants, children with physical and speech difficulties, as well as for exceptionally gifted children.

A monograph published just before the war at Teachers College by Robert A.F. McDonald on the "adjustment of school organization to various population groups" outlined the history of special education. Both the history and the national pattern of special education in 1914 can be seen in Table III, adapted from his study. The figures inside the bars indicate the total number of institutions providing for individual differences in 1914. The figures in parentheses show the date when

instruction first began in a particular type of institution, school, or class.[23]

TABLE III

PROVISIONS FOR INDIVIDUAL DIFFERENCES IN PRIVATE INSTITUTIONS, PUBLIC INSTITUTIONS, AND PUBLIC SCHOOLS, 1914

DEAF	47 Private Institutions (1817) 58 Public Institutions (1823) 73 Cities having day classes (1869)
JUVENILE DELINQUENT	53 Private Institutions (1825) 81 Public Institutions (1826) 158 Cities having day classes (1876)
BLIND	13 Private Institutions (1832) 48 State Institutions (1837) 20 Cities having day classes (1900)
DEPENDENT AND NEGLECTED CHILDREN	66 Private Schools (1729) 20 State Schools (1861)
FEEBLE-MINDED, RETARDED, AND EPILEPTIC	19 Private Institutions (1848) 39 State Institutions (1851) 284 Cities with day classes (1896)
CRIPPLES	30 Private Institutions (1861) 6 Private Day Schools (1890) 6 Public Institutions (1888) 7 Cities having day classes (1899)
NON-ENGLISH-SPEAKING IMMIGRANTS	237 Cities having evening classes (1870) 95 Cities with day classes (1898)
OPEN-AIR SCHOOLS	137 Cities (1908)
SPEECH DEFECTIVES	17 Cities with day classes (1908)
EXCEPTIONALLY GIFTED	22 Cities (1900)

As McDonald noted, special provision for individual differences began first with students who were "subnormal," especially those with "marked pathological defects." His study did not consider school provisions for differences in age, sex, or race. Typically, these provisions were initiated by private or philanthropic agencies. State institutions then took over custody of the child and later city school systems established

special day classes. His study outlined three stages in the "finer adjustments of the public school" to student differences: designation of special teachers, development of new teaching methods, and adaptation of the school curriculum. "In the education of defectives," he observed, "there is a general tendency to introduce vocational guidance, training, and placement studies in the same school." "The mind of the public school profession," he said, "has gradually come to recognize the problem of the varying groups and to perceive the need of special means for school methods."[24]

A survey was conducted for the U.S. Bureau of Education about the same time by Leonard Ayres, James Van Sickle, superintendent of schools in Baltimore, and Lightner Witmer, professor of psychology and director of the Psychological Clinic at the University of Pennsylvania. The study confirmed that provision for individual differences in the schools had become widespread. For the survey some twelve hundred questionnaires were sent to the superintendents of all municipal school systems; approximately 70 percent sponded. The survey showed that provision for "exceptional" children was greatest in the Northeast and Midwest, areas of highest population concentration. According to the study, 10 percent of the schools provided special classes for the "physically exceptional" (blind, deaf, dumb, and crippled, and children with physical and speech defects), and 17 percent provided classes for the "morally exceptional" (delinquents). The widest use of special classes was for the "environmentally exceptional" (non-English speaking immigrants and late-entering students)--39 percent--and the "mentally exceptional" (defectives, backward, and gifted)--42 percent. Evidence that schools were making extensive provision for "defective," "backward," and "gifted" children would

have important consequences for the growth of the testing movement. It was precisely these kinds of students, the "abnormal" students, that the early intelligence tests were designed to identify, and it was for these special classes that the tests were first used. Meanwhile, however, the schools were experimenting with systems for classifying not only "exceptional" individuals but all students, a development that would pave the way for even wider application of intelligence tests.[25]

A parallel effort to provide for individual differences was the development of the age-graded school. In 1856 Henry Barnard, superintendent of schools in Connecticut, pointed out the enormous obstacles preventing effective teaching in the common schools. "There is a large amount of physical suffering and discomfort, as well as great hindrances in the proper arrangement of scholars and classes, caused by crowding the older and younger pupils into the same school-room..." for "there cannot be a regular course of discipline and instruction, adapted to the age and proficiency of pupils," he said. Such a situation taxed the resources of even the best teacher: "With studies ranging from the alphabet and the simplest rudiments of knowledge, to the higher branches of an English education, a variety of methods of instruction and illustration are called for, which are seldom found together, or in an equal degree, in the same teacher." As a solution, Barnard proposed a system of classification by age and attainment. "By arranging scholars of the same general division in different classes, no pupil need be detained by companions who have made, or can make less progress, or be hurried over lessons and subjects in a superficial manner, to accomodate the more rapid advancement of others."[26] Barnard's comments evidently reflected the concerns of school people in many villages and large cities, for by

the 1870s the system of age-grading had become extensive among the larger, urban schools.

When William T. Harris, superintendent of schools in St. Louis in the 1870s, pondered the effects of age-grading on his students, the flaws in age-grading, rigidly applied, quickly became apparent. In his annual reports Harris discussed the varied attainments of his pupils: irregular attendance, different starting ages, and most importantly differences in "temperament and character--the slow and the swift, the weak and the strong, the careless and the earnest, the mature and the immature, the industrious and the indolent." A rigid age-grading system was inadequate to deal with such differences, he said. On the one hand, such a system made the school a "lifeless machine, a Procrustean bed," recalling the Greek giant who stretched or shortened his captives to fit an iron bed. On the other, a rigid age-grading not only held back the bright, but injured the less able. "The pupil who tries his best and then fails is deeply injured, and is apt to endeavor to preserve his self-respect by some sort of subterfuge," he said. The root of all bitterness is loss of self-respect; the man or child who goes about thinking himself shut out from participation in the highest by his own natural incapacity is like one enclosed in a tomb while yet living." To solve these problems, Harris devised a simple solution--flexible-grading. By promoting students frequently at intervals of approximately six weeks, the St. Louis schools allowed students to progress at varying rates.[27]

In the next two decades, the St. Louis system of flexible grading became fairly widespread. A survey reported by the U.S. Commissioner of

Education in 1891 showed that 44 percent of the elementary school systems in cities over 4,000 made individual promotions more frequently than once a year. The general practice was to use flexible promotions in the early grades. Over half the schools reported flexible grading in first grade, while some 17 percent used it in the senior year of high school.[28]

In the last quarter of the nineteenth century, school administrators and teachers invented a panoply of "plans" to provide for individual differences. Some, like the St. Louis system, assumed uniform content but varied the pace of instruction and the amount of individual attention. In the Batavia Plan, as described by Cubberley, teaching assistants aided slow students, while the "left-overs" and "dull or retarded pupils" were lumped in "ungraded classes." Other programs with flexible grading included: the Pueblo Plan (individual and small group instruction with the pupil advancing at any time), the Chicago Plan (entire classes promoted when the group had achieved a certain level of performance), group teaching (instruction by teams of teachers) and departmental teaching (instruction in seventh and eighth grade in the fashion common to high schools). Probably the most widely used flexible grading system, the Elizabeth Plan, employed small group and individual instruction with promotions at any time.[29]

Of a more radical bent were programs which varied the courses of study for students of different abilities. The most widely known of this type of classification system was called the Cambridge Plan, an early form of ability grouping. The plan originated with Charles W. Eliot, president of Harvard. Addressing the annual meeting of the National Education Association in 1892 at Saratoga, New York, Eliot deplored the "undesirable uniformity" produced by rigid age-grading. This system, he

said, "must suppress individual differences instead of developing them, and must leave individual capacities undiscovered and untrained, thus robbing the individual of happiness and serviceableness, and society of the fruits it might have enjoyed from the special endowments of thousands of its members."[30]

Francis Cogswell, superintendent of the Cambridge Schools, quickly developed a parallel course of study, or two-track plan, to remedy the malady which Eliot had diagnosed. After the third grade, students were divided into two groups, "average" and "gifted." Those in the first group completed the subsequent course of study in six years, while the "gifted" group was given an enriched program which it completed in four years. At yearly intervals it was possible for students in one group to switch into the other.[31]

A variant of the Cambridge Plan, developed by Superintendent James Van Sickle for the North Denver schools, provided a similar method to meet the needs of the "average" and "gifted" students. Moving to Baltimore, Van Sickle introduced an important modification to the parallel course of study. In the first six grades students followed three different paths, one directed toward the "minimum essentials," a second toward the "average" group, and a third toward the "superior" group. A new type of institution, the intermediate school, then provided a differentiated course of instruction, with students in the third group preparing for advanced academic work and others taking courses oriented towards business, "household arts" and vocational courses. In rough form Van Sickle had created an early version of a three-track plan with a differentiated course of study.[32]

How widespread were these three methods of classifying students? A 1910 survey of 2,000 superintendents, principals, and teachers conducted by the Joint Committee on School Organization of the New York City Teachers' Associations provided evidence of their use in elementary schools. The following table shows the percentage of the educators surveyed who indicated they had tried a particular plan:

TABLE IV

THE GRADING AND PROMOTION OF PUPILS, 1910

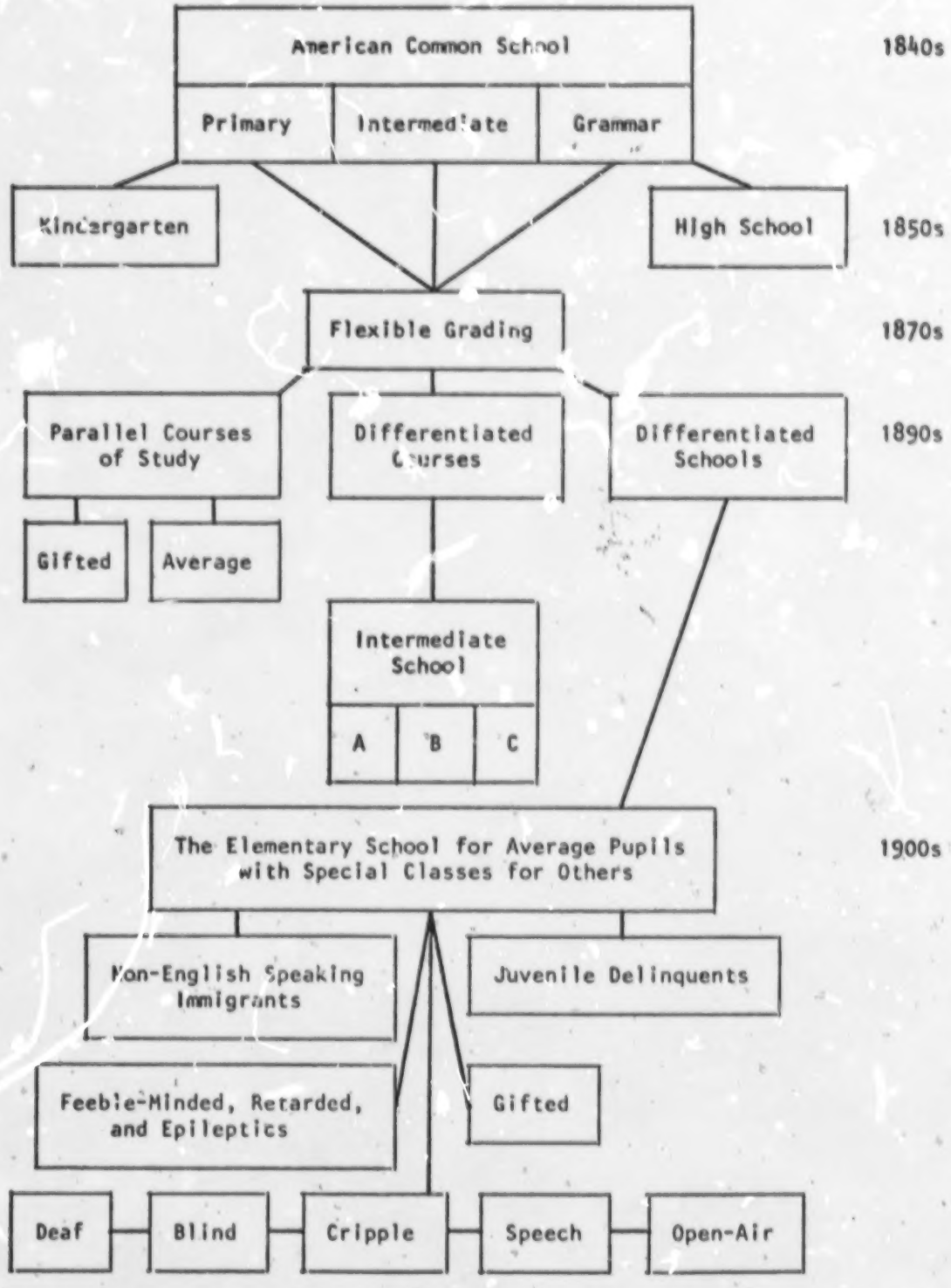
Classification Method	Percent Who Had Tried
Group Teaching	55%
Departmental Teaching	49%
Ungraded Classes	28%
Elizabeth Plan	28%
Pueblo Plan	21%
North Denver Plan	17%
Chicago Plan	16%
Batavia Plan	14%
Cambridge Plan	8%

Though the research methods in this survey were not sophisticated, the results do allow some general observations.[33] It appears that programs involving the least amount of structural reform, such as group teaching and ungraded classes achieved the widest use. Plans such as that used in Cambridge, involving a complex tracking scheme, appear to have been used only infrequently in these early years of the century.[34]

Around the time of the First World War, Cubberley described the structural evolution of public education in the United States. As is shown in Table VII below, the elementary school had made special provisions for individual students and had begun to adopt several plans for differentiated instruction.[35]

TABLE V

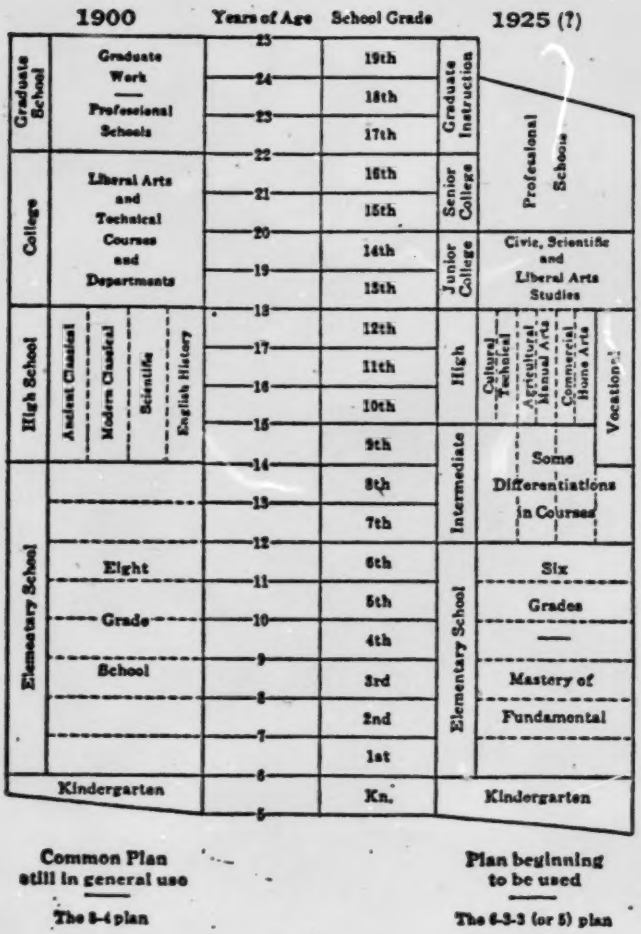
EVOLUTION OF THE STRUCTURE OF ELEMENTARY EDUCATION



A similar evolution in the structure of high schools had led to the creation of four distinct institutions: college preparatory, comprehensive, commercial, and vocational. Cubberley also provided a more detailed description of the internal structure of most urban school systems around the time of the war, as well as his prognosis for further development:[36]

TABLE VI

THE REORGANIZATION OF AMERICAN EDUCATION, 1919



As the structure of American public schools changed significantly between 1840 and 1917, so too did their function as perceived by school administrators and university professors. The transformation of the high school around the turn of the century revealed the most profound shift in the function of education. In 1890 only 6.7 percent of those aged fourteen to seventeen were enrolled in public and private high schools. Already, however, the high school curriculum was changing to meet the needs of a more diversified clientele. In 1892 the National Education Association appointed Harvard President Charles W. Eliot and a group of educators to evaluate high school programs and their relation to college admissions requirements. The Committee of Ten, as it came to be known, recommended four different courses of study--Classical, Latin-Scientific, Modern Languages, and English. Importantly, they concluded that subjects should not be "treated differently for pupils who are going to college, for those who are going to a scientific school, and for those who, presumably, are going to neither." Though the work of the Committee did signal a new direction for the high school, the reforms they proposed did not address the circumstances of the great majority of youth who went directly to work. [37]

How different was the position taken twenty-five years later by the Commission on the Reorganization of Secondary Education in 1917. By then high schools enrolled nearly one-third of the population between fourteen and seventeen. Noting that "within the past few decades changes have taken place in American life profoundly affecting the activities of the individual," the Commission identified seven "cardinal principles" of education: health, command of fundamental processes, worthy home-membership, vocation, citizenship, worthy use of leisure, and

ethical character. From the Commission's view, an exclusive secondary school could no longer fill the needs of an educational system that was quickly becoming universal. The report on "cardinal principles," though obviously not a measure of actual practice, suggested that the educational system was on its way to becoming highly differentiated in function as well as structure.[38]

The Adoption of Tests

America's entrance into World War I in 1917 proved to be the catalyst that fused the interests of applied psychologists who had developed intelligence tests and the needs of administrators who were struggling to adapt the schools to the conditions of the twentieth century. Many psychologists, like Terman, answered the call of the American Psychological Association to join an experiment using intelligence tests in the army. In a story now familiar, the psychologists modified individual intelligence tests and used the new group tests to classify 1.7 million recruits between September 1917 and January 1919. America's involvement in World War I made three important contributions to the national testing movement. In addition to the dramatic classification experiment, the war also shaped the network of professional psychologists and intelligence testers by providing a forum for university professors to share ideas they had been developing and by giving a new group of men skill in using tests. Demobilization led directly to the creation of group intelligence tests for use in the public schools.[39]

Shortly after the war's end, Terman observed that the Army tests "demonstrated beyond question that the methods of mental measurement are capable of making a contribution of great value to army efficiency....That their universal use in the schoolroom is necessary to educational efficiency will doubtless soon be accepted as a matter of course." Terman's prophecy

turned out to be accurate, for in the next decade intelligence testing spread throughout the public school system with astonishing swiftness. Under the auspices of the National Academy of Sciences, Terman helped transform the Army tests into the National Intelligence Tests for school children in 1919, and within a year over 400,000 copies had been sold.[40]

By the mid-1920s testing and classification by ability reached countless schools systems across the country. Psychologists had developed over seventy-five tests of general mental ability and, according to some estimates yearly intelligence test sales approached four million. Thousands of people--university professors, school administrators, teachers, publishers, and the directors of philanthropic foundations--collaborated to introduce intelligence tests into the nation's schools. Although the movement was diffuse, the broad outlines of the impact of testing in the schools can be sketched by using several national surveys conducted at the time.

A survey conducted in 1925 by the U.S. Bureau of Education, "Uses of Intelligence and Achievement Tests in 215 Cities," offered convincing evidence of the rapid increase in the use of tests for classifying students.[41] W.S. Deffenbaugh, chief of the Bureau's City School Division, asked all superintendents in cities over 10,000 to report on the purposes for which they used intelligence and achievement tests in the elementary and secondary public schools. He presented the replies from 215 cities (a 35 percent reponse rate).

Deffenbaugh reported first the use of group intelligence tests:

TABLE VII

THE USE OF GROUP INTELLIGENCE TESTS IN ELEMENTARY, JUNIOR HIGH, AND HIGH SCHOOLS IN CITIES OVER 10,000
1925

Purposes for which tests are used	Elementary schools		Junior high schools		High schools	
	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose
Classification of pupils into homogeneous groups	64	1	56	1	41	1
Supplementing teachers' estimates of pupils' ability	62	2	44	2	33	2
Diagnosis of cause of failure	46	3	29	3	24	3
Establishment of classes for subnormal children	43	4	14	10	7	13
Extra promotions	40	5	21	4	8	11
Comparison with other school systems	26	6	18	7	13	7
Admission to first grade of elementary school	25	7	0	21	0	21
Placement of new pupils from other schools	23	8	10	6	10	10
Regular promotion of pupils	22	9	15	9	6	16
Determining comparative efficiency of teachers	20	10	11	13	10	9
Establishment of classes for supernormal children	20	11	6	17	2	20
Diagnosis of cause of success	19	12	16	6	12	8
Demotions	17	13	8	16	7	12
Determining changes in method of presentation of lessons	14	14	13	11	8	15
Determining changes in subject matter of courses of study	11	15	9	14	7	14
Determining class marks	10	16	7	19	4	18
Establishing special supervised study groups	8	17	6	17	3	19
Vocational guidance	0	-------	13	12	17	6
Determining number of courses to be carried at one time by high school pupils	0	-------	0	14	21	5
Guidance in the selection of high school course	0	-------	10	5	24	4
Admission to organized school activities	0	-------	3	20	5	17

The data showed elementary schools were most likely to use tests, high schools least likely, with junior highs in between. Group tests were used most often to classify pupils into homogeneous groups. Nearly two-thirds of the elementary schools, over one-half of the junior high schools, and two-fifths of the high schools employed the tests for this purpose. Supplementing teachers' estimates of student ability and diagnosing failure were the next most frequent use in all schools. Junior high schools and especially high schools reported extensive use of tests in educational and vocational guidance. There was variation among the three types of schools in the other uses of group tests, which included establishing special classes and making rapid promotions.

The survey then described the use of individual intelligence tests:

TABLE VIII

THE USE OF INDIVIDUAL INTELLIGENCE TESTS IN ELEMENTARY, JUNIOR HIGH, AND HIGH SCHOOLS IN CITIES OVER 10,000
1925

Purposes for which tests are used	Elementary schools		Junior high schools		High schools	
	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose
Establishment of classes for subnormal children	67	1	18	3	8	7
Classification of pupils into homogeneous groups	40	2	22	2	14	3
Supplementing teachers' estimates of pupils' ability	38	3	24	1	16	2
Diagnosis of cause of failure	37	4	23	4	25	1
Extra promotions	33	5	16	5	7	10
Demotions	25	6	9	7	4	16
Placement of new pupils from other schools	23	7	9	8	3	12
Diagnosis of cause of success	16	8	12	6	11	4
Regular promotion of pupils	13	9	7	10	2	19
Establishment of classes for supernormal children	13	10	8	9	6	15
Admission to first grade of elementary school	12	11				
Determining comparative efficiency of teachers	9	12	6	13	7	8
Comparison with other school systems	8	13	8	14	7	9
Determining changes in method of presentation of lessons	8	14	5	15	5	13
Determining changes in subject matter of courses of study	7	15	4	17	5	14
Establishing special supervised study groups	5	16	4	18	1	20
Determining class marks	4	17	3	19	3	17
Vocational guidance			7	11	10	6
Guidance in selection of high-school course			7	12	10	5
Adjusting to organized school activities			2	20	5	18
Determining number of courses to be carried at one time by high-school pupils			4	18	7	11

All schools used individual tests less frequently than group tests. Establishing classes for subnormal children, classifying students into homogeneous groups, supplementing teachers' judgments, and diagnosing failure ranked among the most frequent uses. Use was most widespread among elementary schools, where individual and group intelligence tests served almost identical purposes.

Finally, Deffenbaugh reported the use of standardized educational, or achievement, tests:

TABLE IX

THE USE OF STANDARDIZED EDUCATIONAL TESTS IN ELEMENTARY, JUNIOR HIGH, AND HIGH SCHOOLS IN CITIES OVER 10,000
1925

Purposes for which tests are used	Elementary schools		Junior high schools		High schools	
	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose	Per cent of cities	Rank of purpose
Supplementing teachers' estimates of pupils' ability	61	1	41	1	27	1
Comparison with other school systems	57	2	34	3	16	3
Classification of pupils into homogeneous groups	54	3	35	2	27	2
Diagnosis of cause of failure	44	4	28	4	14	5
Extra promotions	41	5	20	5	0	8
Determining comparative efficiency of teachers	38	6	18	7	13	4
Regular promotion of pupils	32	7	21	6	7	11
Placement of new pupils from other schools	27	8	15	9	6	16
Establishment of classes for subnormal children	25	9	9	15	2	18
Determining changes in method of presentation of lessons	20	10	15	8	8	10
Determining changes in subject matter of courses of study	20	11	14	11	7	13
Demotions	19	12	8	16	3	17
Diagnosis of cause of success	16	13	15	10	6	15
Determining class marks	14	14	10	14	6	14
Establishment of classes for supernormal children	11	15	6	17	1	20
Establishing special supervised study groups	10	16	8	16	2	18
Guidance in the selection of high-school courses	.....	.....	12	12	10	6
Vocational guidance	.....	.....	11	13	9	7
Admission to organized school activities	.....	.....	3	20	7	11
Determining number of courses to be carried at one time by high-school pupils	.....	.....	4	19	9	7

Standardized achievement tests were about as widely used as group intelligence tests in all schools. They were used for different purposes, though, especially to evaluate the performance of schools. Achievement tests were used to classify students into homogeneous groups, but all schools were more likely to use group intelligence tests to that end, especially in junior highs and high schools.

In a follow-up analysis of the Bureau of Education survey, Bertha Hebb described "The Classification of Elementary School Pupils into Homogeneous Groups." She reported that 80 percent of the cities (172 of 215) were using standardized tests to classify elementary school pupils. Analyzing the data by size of city, she found little variation. To place students schools were most likely to use group intelligence tests,

then standardized achievement tests, and finally individual intelligence tests.[42]

In 1926 the Bureau of Education released another study, "Cities Reporting the Use of Homogeneous Grouping."[43] This survey, also sent to superintendents in cities of 10,000 or more, probed similar issues regarding intelligence testing and student classification. A tabulation of the returns from 292 cities (probably a better response rate even though the number of cities had increased over the previous year) showed the following pattern in the provision of ability groups:

TABLE X

HOMOGENEOUS GROUPING IN CITIES OVER 10,000
1926

Size of City	Elementary Schools			Junior High Schools			High Schools		
	Number with Ability Groups	N	%	Number with Ability Groups	N	%	Number with Ability Groups	N	%
100,000+	36	40	90.0	28	40	70.0	26	40	65.0
30,000-100,000	66	89	74.2	57	89	64.0	36	89	40.4
10,000-30,000	145	163	89.0	119	163	73.0	81	163	49.7
TOTAL	247	292	84.6	204	292	69.9	143	292	49.0

As the survey revealed, ability grouping had become very widespread by the middle of the 1920s, especially in cities over 100,000. Elementary schools were the most likely to have ability grouping (85 percent), followed by junior high schools (70 percent) and high schools (49 percent).

The survey also sought to determine the methods used to group students, reporting the following instances in which IQ or mental age was employed:

TABLE XI

THE USE OF INTELLIGENCE TESTS IN CLASSIFICATION
1926

Size of City	No. of Cities	Total Cities	Percent
100,000+	36	40	90.0
30,000-100,000	67	89	75.3
10,000-30,000	147	163	90.2
TOTAL	250	292	85.6

Thus by the 1920s school systems across the country had become highly differentiated and frequently used intelligence and achievement tests to classify students.[44]

Perspectives on Reform

In the half century after the progressive period, the use of intelligence tests for ability grouping became standard practice throughout American education. As this essay has shown, the adoption of intelligence tests by the schools was a complex development. Certainly tests were adopted as part of the reform program fashioned by the network of applied psychologists and school people. While the network itself often viewed testing as a means to improve the schools and society, however, not everyone shared their vision, especially recent immigrants and blacks. In spite of this disagreement, the adoption of tests proceeded quickly for several reasons. The evolution of the schools around the turn of the century into more complex institutions that provided a variety of curricular paths created a need for an efficient means of sorting students. In the early years of the twentieth century, educators used students' academic performance, teachers' judgments, and, increasingly, standardized achievement tests to accomplish this task. The building intelligence testing

movement was then quickened by the catalyst of the First World War. Those who argue that intelligence tests were merely imposed on the schools by an elite have a point, but they fail to acknowledge the consensus among school people that the use of tests promoted a more efficient and scientific education.

In the last two decades the progressive reform of intelligence testing and ability grouping has come under sharp attack. Revisionist historians and social activists have criticized intelligence tests as racially biased, elitist, and as obstacles to social equality. Within the past few years the entire testing industry--especially standardized aptitude tests used for college admission--has become the focus of national concern. Recently two states passed laws regulating the use of tests, while some twelve other states and the United States Congress have legislation pending. Ralph Nader has joined the fray by releasing a highly critical study which charges that aptitude tests "were conceived by the upper class for the upper class and have served as a formidable barrier to millions of students, unjustly diminishing their higher education and career opportunities."[45] Historically it cannot be denied that the rhetoric and the programs of progressive reformers occasionally led to inequality in the schools. Yet as the nation seeks to "reform a reform" and to reevaluate policy regarding the use of tests, it is essential to consider the origins of intelligence testing in the progressive era. In that history there are no simple explanations about the adoption of intelligence tests; similarly, there are no simple solutions for developing a fair policy regarding the use of tests today.

NOTES

1. Lewis M. Terman, "The Use of Intelligence Tests in the Grading of School Children," Journal of Educational Research 1 (Jan. 1920): 20.

2. For some recent discussions of the history of intelligence testing, see the following: Clarence Karier, "Testing for Order and Control in the Liberal Corporate State," Educational Theory 22 (Spring 1972); Russell Marks, "Testers, Trackers, and Trustees: The Ideology of the Intelligence Testing Movement in America 1900-1954" (Ph.D. dissertation, University of Illinois, 1972); Joel Spring, "Psychologists and the War: The Meaning of Intelligence in the Alpha and Beta Tests," History of Education Quarterly 12 (Spring 1972): 13; Samuel Bowles and Herbert Gintis, Schooling in Capitalist America (New York: Basic Books, 1976); David K. Cohen and Marvin Lazerson, "Education and the Corporate Order," Socialist Revolution 2 (April 1972); and Leon J. Kamin, The Science and Politics of I.Q. (Potomac, Md.: Lawrence Erlbaum Assoc., 1974).

3. Paul Davis Chapman, "Schools as Sorters: Testing and Tracking in California, 1910-1925," Paper Presented at AERA Annual Meeting, April 8, 1979; see also Chapman, "Schools as Sorters: Lewis M. Terman and the Intelligence Testing Movement, 1890-1930" (Ph.D. dissertation, Stanford University, 1980) upon which this paper is based.

4. Kimball Young, "The History of Mental Tests," Pedagogical Seminary 31 (1924), 2.

5. Edwin G. Boring, A History of Experimental Psychology, 2nd ed. (New York: Appleton-Century-Crofts, 1950).

6. For an informative study of Binet, see Theta H. Wolf, Alfred Binet (Chicago: University of Chicago Press, 1973).

7. Lewis M. Terman, "The Binet-Simon Scale for Measuring Intelligence," Psychological Clinic 5 (Dec. 1911): 204, 205.

8. Lewis M. Terman, Measurement of Intelligence (Boston: Houghton Mifflin, 1916); for an example of the critics' view, see Leonard P. Ayres, "The Binet-Simon Measuring Scale for Intelligence: Some Criticisms and Suggestions," Psychological Clinic 5 (1911): 187-96.

9. Lewis M. Terman, "To Help Backward School Children," Stanford Alumnus (Apr. 1914), 1, 4.

10. Lewis M. Terman and Ernest B. Hoag, Health Work in the Schools (Boston: Houghton Mifflin Co., 1914); Terman, The Hygiene of the School Child (Boston: Houghton Mifflin, 1914); and Terman, et al., Surveys in Mental Deviation in Prisons, Public Schools, and Orphanages in California (Sacramento, Calif.: State Board of Charities and Corrections, 1918).

11. Terman, Hygiene, 1.

12. Terman, Surveys, 26; Terman, "San Luis Obispo," (1914): 132; Terman, "Backward Children," Stanford Alumnus, 2.

13. For the early measurement movement, see Joseph M. Rice, The Public School System of the United States (New York: Century, 1893); "A New Basis in Education," and "Obstacles to Rational Educational Reform," The Forum (Dec. 1896 and Jan.-Mar. 1904), reprinted in Scientific Management in Education (New York: Hinds, Noble, and Eldredge, 1913).

14. see Geraldine Joncich, Edward L. Thorndike: The Sane Positivist (Middletown, Conn.: Wesleyan University Press, 1968),

15. Walter S. Monroe, "Existing Tests and Standards," Seventeenth Yearbook of the National Society for the Study of Education (1918), pt. 2, 17, 161-90; Monroe, Educational Tests and Measurements (Boston: Houghton Mifflin, 1917).

16. Jesse B. Sears, The School Survey: A Textbook on the Use of School Surveying in the Administration of Public Schooling (New York: Houghton Mifflin, 1925).

17. Eugene A. Kilfenecker, "Bureaus of Research in City School Systems," Seventeenth Yearbook of the National Society for the Study of Education (1918), pt. 2, 52.

18. J. E. W. Wallin, The Mental Health of the School Child (New Haven: Yale University Press, 1914).

19. Ibid., 393.

20. Ibid., 23, 138, 152, 154.

21. Ibid., 23.

22. On the evolution of school structure, see Tyack, One Best System and Samuel P. Hayes, "The Politics of Reform in Municipal Government in the Progressive Era," Pacific Northwest Quarterly 55 (Oct. 1964); 157-69.

23. Robert A. F. McDonald, Adjustment of School Organization to Various Population Groups, Contributions to Education, no. 75 (New York: Teachers College, Columbia University, 1915), 108, 120-45. Even though McDonald's study did not define the sample, the data nonetheless seem useful in suggesting how provisions were made for special education.

24. Ibid., 104, 106, 105.

25. J. H. Van Sickle, Lightner Witmer, and Leonard P. Ayres, "Provisions for Exceptional Children in City School Systems," Department of the Interior, Bureau of Education, Bulletin no. 14 (1911): 5-92.

26. Henry P. Barnard, "Gradation of Public Schools, with Special Reference to Cities and Large Villages," American Journal of Education 2 (Dec. 1856): 455, 456, 458.

27. William T. Harris, "The Development of the Short-Interval System in St. Louis," in "Classification and Promotion of Pupils," Report of the United States Commissioner of Education, vol. 1, 1898-99, 99, 328, 312, 304.

28. J. C. Boykin, "Class intervals in City Public Schools," Report of the United States Commissioner of Education, vol. 2, 1890-91, 981-1009. A questionnaire regarding grading practices in schools was addressed to superintendents in all cities and villages of 4,000 and over. Of 816 questionnaires sent approximately 465 were returned, a response rate of 57 percent. Note that column entitled "Total Replies" does not equal 465 since for each grade approximately 25 answers were labeled "vague or indefinite."

29. Ellwood P. Cubberley, Public Education in the United States: A Study and Interpretation of American Educational History (Boston: Houghton Mifflin, 1916), 314.

30. Charles W. Eliot, "Undesirable and Desirable Uniformity in Schools," Educational Reform (New York: Century, 1898), 271-300.

31. Francis Cogswell, "Promotions in the Grammar Schools of Cambridge, Mass.," in "Classification and Promotion of Pupils," Report of the United States Commissioner of Education, vol. 1, 1898-99, 346-48.

32. Ellwood P. Cubberley, Public School Administration (Boston: Houghton Mifflin, 1916), 306.

33. C. S. Hartwell, "The Grading and Promotion of Pupils," Addresses and Proceedings of the National Education Association 48 (1910): 296; replies were received from half the group surveyed. Responses came from 325 superintendents; 513 principals; 20 teachers; 25 college presidents and professors; 66 normal school presidents, teachers, and supervisors of training schools; and 21 others including members of boards of education. Responses came from all 48 states and territories, Puerto Rico, Manila, and Canada. After eliminating "vague replies," 965 answers were left. Because the survey does not indicate any more precisely how the sample was selected, it is difficult to know the representativeness of the data.

34. For other discussions of pupil classification during this period, see: S. T. Dutton and David Snedden, "Grading and Promotion," The Administration of Public Education in the United States (New York: Macmillan Co., 1908); H. S. Jones, et al., "Report of the Committee on City School Systems," Addresses and Proceedings of the National Education Association 26 (1886): 276-84; William H. Holmes, School Organization and the Individual Child (Worcester, Mass.: Davis Press, 1912); Julia Richman, "A Successful Experiment in Promoting Pupils," Education Review 18 (June 1899): 23-29; William J. Shearer, The Grading of Schools (New York: H. P. Smith, 1898); and James H. Van Sickle, "Plan of the North-Side Schools of Denver," in "Classification and Promotion of Pupils," Report of the United States Commissioner of Education, vol. 1, 1898-99, 341-46.

35. Adapted from Cubberley, Public Education, 401, 429.

36. Ibid., 460.

37. W. Vance Grant and C. George Lind, Digest of Education Statistics, 1977-78 (Washington, D.C.: Government Printing Office, 1978), 44; Committee of Ten, Report (New York: Published for the National Education Association by American Book Co., 1894), 6; for discussions of the Cardinal Principles, see Krug, Sizer, and Clarence Karier, Man, Society, and Education: A History of American Educational Ideas (Glenview, Ill.: Scott, Foresman and Co., 1967), 69-79.

38. Digest, 44; Cardinal Principles of Secondary Education: A Report of the Commission on the Reorganization of Secondary Education, Appointed by the National Education Association, by Clarence D. Kingsley, Chairman, Department of the Interior, Bureau of Education, Bulletin no. 35, 1918 (Washington, D.C.: Government Printing Office, 1918), 7, 11.

39. For an excellent discussion of the use of tests in World War I, see Daniel J. Kevles, "Testing the Army's Intelligence: Psychologists and the Military in World War I," Journal of American History 55 (Dec. 1968): 565-581.

40. Lewis M. Termar, The Intelligence of School Children (Boston: Houghton Mifflin, 1919), xiv.

41. W.S. Deffenbaugh, "Uses of Intelligence and Achievement Tests in 215 Cities," Department of the Interior, Bureau of Education, City School Leaflet no. 20 (Washington D.C., 1925).

42. Bertha Y. Hebb, "The Classification of Elementary School Pupils into Homogeneous Groups," American School Board Journal 72 (May 1926): 52, 140.

43. "Cities Reporting the Use of Homogeneous Grouping and of the Winnetka Technique and the Dalton Plan," Department of the Interior, Bureau of Education, City School Leaflet no. 22 (Washington, D.C., 1926).

44. Ibid.

45. The College Board, "Updated Status Report on State and National Legislation Regulating Standardized Admissions Tests," February 1980; The College Board, "Memo to the Members," February 1980.